

Chernoff's Theorem and Discrete Time Approximations of Brownian Motion on Manifolds

Oleg G. Smolyanov

Faculty of Mechanics and Mathematics, Moscow State University, Russia

Heinrich v. Weizsäcker

Fachbereich Mathematik, Technische Universität Kaiserslautern, Germany

Olaf Wittich

Mathematisches Institut, Universität Tübingen, Germany

Abstract. Let $(S(t))_{t \geq 0}$ be a one-parameter family of positive integral operators on a locally compact space L . For a possibly non-uniform partition of $[0, 1]$ define a finite measure on the path space $C_L[0, 1]$ by using a) $S(\Delta t)$ for the transition between any two consecutive partition times of distance Δt and b) a suitable continuous interpolation scheme (e.g. Brownian bridges or geodesics). If necessary normalize the result to get a probability measure. We prove a version of Chernoff's theorem of semigroup theory and tightness results which yield convergence in law of such measures as the partition gets finer. In particular let L be a closed smooth submanifold without boundary of a manifold M . We prove convergence of Brownian motion on M , conditioned to visit L at all partition times, to a process on L whose law has a density with respect to Brownian motion on L which contains scalar, mean and sectional curvatures terms. Various approximation schemes for Brownian motion on L are also given.

Keywords: Approximation of Feller semigroups, geodesic interpolation, Brownian bridge, (mean, scalar, sectional) curvature, Wick's formula, pseudo-Gaussian kernels, conditional process, infinite dimensional surface measure



© 2008 Kluwer Academic Publishers. Printed in the Netherlands.

Table of Contents

1	Introduction	3
2	Chernoff's Theorem	5
	2.1 Chernoff Equivalence	6
3	The Pinning Construction	8
	3.1 Discrete Time Approximations of Feller processes	8
	3.2 Continuous Interpolations	12
4	Gaussian Integrals	16
	4.1 Wick's Formula	17
5	Proper Families Equivalent to the Heat Semigroup on Manifolds	19
	5.1 Introduction	19
	5.2 First Examples	20
	5.3 The Heat Equation on a submanifold via the restriction of a pseudo-Gaussian kernel	23
	5.4 Proof of Theorem 4	24
	5.5 Pinning the Heat Kernel	29
6	Limit Densities for the Pinning Construction under global normalization	31
	6.1 Introduction	31
	6.2 The Pinning Construction as Conditional Probability	33
7	Conclusion	34

1. Introduction

This paper is an extension of earlier work published in two conference proceedings ((Smolyanov, v. Weizsäcker and Wittich, 2000) and (Smolyanov, v. Weizsäcker and Wittich, 2003). The classical Chernoff Theorem states roughly that in the strong sense $S(t/r)^r \rightarrow e^{t\bar{D}S}$, where $S = (S(t))_{t \geq 0}$ is a strongly continuous operator family on a Banach space and $\bar{D}S$ its derivative at $t = 0$, cf. (Ethier and Kurtz, 1986), p. 32, Cor. 6.6. This means, under some technical assumptions, that two operator families S, S' with $DS = DS'$ yield in the limit the same semigroup. We call such families *Chernoff equivalent*. We are particularly interested in the case of families of positive integral operators on a smooth closed manifold L without boundary. The iterations $S(t/r)^r f$ are then given by iterated integrals with finite and positive kernels. Slightly extending a result from (Smolyanov, v. Weizsäcker and Wittich, 2000) we actually give a version of Chernoff's Theorem (Proposition 1) for nonuniform partitions \mathcal{P} of the time interval $[0, t]$.

For every starting point $x \in L$, time horizon $t > 0$ and partition $\mathcal{P} = 0 = t_0 < \dots < t_r = t$, the finite family of operators $S(\Delta t_k) \circ \dots \circ S(\Delta t_1)$, $0 \leq k \leq r$ defines a finite measure $\mathbb{P}_{\mathcal{P}}^x$ on $L^{\mathcal{P}}$. It can then be extended to a measure on the path space $C_L[0, t]$ by continuous interpolation, either deterministically or using suitable conditional distributions, or more simply on the path space $D_L[0, t]$ if we extend the discrete paths as stepfunctions.

This construction, described in detail in Section 3, depends on the family S and on the interpolating measures. It will be called the *pinning construction*. We are interested in the possible weak limit of these measures as the mesh of the partition tends to 0. The Chernoff result implies convergence of the finite dimensional marginals.

If the family S is Chernoff equivalent to a Feller semigroup we prove in section 3.1, in extension of similar results in (Ethier and Kurtz, 1986), the tightness of the resulting step processes over the path space $D_L[0, t]$. If the limit process has continuous paths and interpolation by geodesics is used then even tightness and hence convergence in law over $C_L[0, 1]$ follows. Based on a Large-Deviation result from (Wittich, 2005), we allow also interpolation by Brownian bridges. If $L \subset M$ is isometrically embedded into another Riemannian manifold M , we may even interpolate by Brownian bridges in the ambient manifold. (Theorem 3).

A particular feature of our results is the fact that the positive operators S_t need not to be normalized in the sense that the associated measures $q(t, x, -)$ are allowed to have finite total mass $\neq 1$. For example if their densities are the restrictions of probability densities on the larger manifold M to L , then the measures $\mathbb{P}_{\mathcal{P}}^x$ constructed above are not probability measures.

There are two different normalization procedures. First one can pass at the beginning from S_t to the associated probability kernels by normalizing each $q(t, x, -)$. This gives a family \tilde{S} which may be Chernoff equivalent to a Markov semigroup. In section 5 we give a couple of examples which are all equivalent to the heat semigroup on L . In order to verify this equivalence we use a detailed study of the short time behaviour of Gaussian integrals from (Smolyanov, v. Weizsäcker and Wittich, 2003) which we review in section 4. After these preparations the remaining work lies in the local differential geometry which one needs for the Taylor expansions of the normalization coefficients.

The second possibility is to renormalize the measures $\mathbb{P}_{\mathcal{P}}^x$ *after* their construction. This is the content of section 6. The corresponding operator families are no longer Chernoff equivalent to the above Markov semigroup but nevertheless the resulting limit measures are equivalent to the law of the Markov process obtained by the first normalization procedure. In all situations studied in section 5 one can even calculate the Radon-Nikodym density with respect to Wiener measure over $C_L[0, 1]$ explicitly by a combination of curvature terms. This includes as a special case a new proof of a result of (Andersson and Driver, 1999). However for us the most important example is the Brownian motion on a larger manifold M which is conditioned to visit the embedded manifold L at all partition times. In this case we get the following main result of this paper.

THEOREM 1. *Let $x \in L$ and $L(\varepsilon) := \{x \in M : d_M(x, L) < \varepsilon\}$ the tubular ε -neighborhood of L and \mathcal{P}_k be a sequence of partitions of the unit interval with mesh $|\mathcal{P}_k| \rightarrow 0$ as $k \rightarrow \infty$. Then the limit law of the conditional Brownian motions on M*

$$\lim_{k \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \mathbb{W}_M^x(d\omega | \omega(t_i) \in L(\varepsilon), t_i \in \mathcal{P}_k) = \mu_L^x \quad (1)$$

exists and is equivalent to the Wiener measure \mathbb{W}_L^x on the submanifold with density

$$\frac{d\mu_L^x}{d\mathbb{W}_L^x}(\omega) = c \cdot e^{\int_0^1 (\frac{1}{4}\text{Scal}_L - \frac{1}{8}|\tau_\phi|^2 - \frac{1}{12}(\overline{R}_{M/L} + \overline{\text{Ric}}_{M/L} + \text{Scal}_M))(\omega(s))ds} \quad (2)$$

where the constant c normalizes to total mass 1.

Remark. This limit law can be interpreted as the infinite dimensional surface measure which is induced by \mathbb{W}_M^x on the submanifold $C_L[0, 1]$ of the path space $C_M[0, 1]$. This interpretation is underlined by the fact that the two limits in (1) may be interchanged without affecting the result. The proof of this fact requires however very different techniques and is even more involved. It has been established in (Sidorova, Smolyanov, v. Weizsäcker and Wittich, 2004) for $M = \mathbb{R}^m$ with tools from stochastic analysis. In this case (Sidorova, 2004) provides an explanation of this interchangeability of the two limiting procedures which is independent of our present methods. The general case will be treated with perturbation theoretical methods in (Wittich, 2005) and (Sidorova, Smolyanov, v. Weizsäcker and Wittich, 2005).

2. Chernoff's Theorem

In this section we formulate a slightly extended form of the Chernoff type result from (Smolyanov, v. Weizsäcker and Wittich, 2000). In contrast to the usual versions we consider convergence along nonuniform partitions of the time parameter. In (Smolyanov, v. Weizsäcker and Wittich, 2000) we treated only contractions whereas in the result below more general bounded operators are allowed.

DEFINITION 1. Let $B(V)$ denote the space of bounded linear operators on the Banach space V . A strongly continuous family $S : [0, \infty) \rightarrow B(V)$ with $S(0) = 1$ is called **proper**, if

$$\|S(t)\| = 1 + O(t) \quad (3)$$

as $t \downarrow 0$, and if there is an operator $(A, \mathcal{D}(A))$ which is the generator of a strongly continuous semigroup (e^{tA}) on V such that

$$\frac{S(t) - I}{t} f \rightarrow Af \quad (4)$$

as $t \downarrow 0$, for all $f \in V$ of the form $f = e^{aA}g$ with $a > 0$ and $g \in V$. The operator $(A, \mathcal{D}(A))$ will be also denoted by DS .

Remarks 1. Clearly the operator A is uniquely determined by the family $S(t)$. Therefore the notation DS is justified. However many proper families may lead to the same operator A . This will be discussed in the next subsection.

2. The vectors of the form $e^{aA}g$ form a core for A , i.e. they are dense in

the domain $\mathcal{D}(A)$ with respect to the graph norm. We do not require the convergence (4) for *all* $f \in \mathcal{D}(A)$ since we do not need it in the following proof and our formally weaker assumption is easier to verify in the situations to be studied later. However in contrast to the usual form of Chernoff's Theorem we do not know whether it suffices for the following result to require the above convergence for $f \in \mathcal{D}$ where \mathcal{D} is an *arbitrary* core of the operator A . Note that the operators $(S(t) - I)/t$ may fail to be uniformly bounded in the graph norm of $\mathcal{D}(A)$. Note also in the usual Chernoff expression $S(t/r)^r$ the factors commute whereas in general in a product of the form $S(t_1) \cdots S(t_r)$ they do not. Therefore it is not surprising that in comparison with the usual statement we need slightly stronger assumptions.

PROPOSITION 1. *Let $S(t)$, $t \geq 0$ be a proper family of linear operators on V and let $A = DS$. Let $(t_i^n)_{1 \leq i \leq r_n}$, $n \in \mathbb{N}$, be positive numbers such that $\sum_{i=1}^{r_n} t_i^n \rightarrow t \geq 0$ and $\max_i t_i^n \rightarrow 0$. Then, for all $f \in V$, we have*

$$S(t_1^n) \cdots S(t_{r_n}^n) f \rightarrow e^{tA} f$$

as $n \rightarrow \infty$.

Proof: In the case of contractions this is Proposition 3 of (Smolyanov, v. Weizsäcker and Wittich, 2000). The same proof works under the weaker assumption (3) since this condition implies the existence of a number $q > 0$ such that $\|S(t)\| \leq e^{qt}$ for all sufficiently small t and thus

$$\|S(t_1) \cdots S(t_r)\| \leq e^{q \sum_{i=1}^r t_i} \quad (5)$$

as soon as $\max_i t_i$ is small enough. \blacksquare

The usual application of Chernoff's Theorem is the construction of discrete *semigroups* $U^{(n)}$, $n = 1, 2, \dots$ approximating a strongly continuous semigroup e^{tA} in the sense of (Kato, 1980), (3.9), p. 511. The key observation that the limit semigroup only depends on the derivative of the family S at $t = 0$, stays valid in our setting. This will be formalized in the next section.

2.1. CHERNOFF EQUIVALENCE

Let Π denote the set of all proper families. We focus on the map which assigns to each proper family its corresponding contraction semigroup. The map $P : \Pi \rightarrow \Pi$ given by

$$P(S)_t := e^{tDS}$$

we call *Chernoff map*. P maps proper families onto the subset $\Sigma \subset \Pi$ of strongly continuous contraction semigroups. Σ remains pointwise fixed under P . We are interested in the attracting domains for each fixpoint, i.e. the set of preimages of a given semigroup.

DEFINITION 2. *Two proper families $S, T \in \Pi$ are called Chernoff equivalent if one of the two equivalent conditions holds*

$$(i) \quad P(S) = P(T),$$

$$(ii) \quad DS = DT.$$

The following simple criteria for Chernoff equivalence will be applied in the sequel.

LEMMA 1. *Let S with $DS = (A, \mathcal{D}A)$. Let $T = (T(t))$ be a family of operators satisfying the bound (3) and*

$$\|T(t)f - S(t)f\| = o(t) \quad (6)$$

for $t \downarrow 0$ and all f of the form $f = e^{aA}g$, $a > 0$, $g \in V$, then T is also proper and Chernoff equivalent to S .

Proof: Let $f = e^{aA}g$, $g \in V$ be given. Then by our assumption (6) and by (4) we get

$$\lim_{t \rightarrow 0} \frac{T(t)f - f}{t} = \lim_{t \rightarrow 0} \frac{S(t)f - f}{t} = Af.$$

Thus by definition 1 $DT = (A, \mathcal{D}(A))$, i.e. T is Chernoff equivalent to S . \blacksquare

LEMMA 2. *Let S be a proper family. Let $c = (c(t))$ be a family of operators satisfying $\|c(t) - I\| = o(t)$. If $T(t) = c(t)S(t)$ then T and S are Chernoff equivalent.*

Proof: This follows from Lemma 1 and the estimate

$$\|T(t) - S(t)\| = \|(c(t) - I)S(t)\| = o(t)\|S(t)\| = o(t).$$

\blacksquare

3. The Pinning Construction

We want to use the results of the preceding section to construct resp. approximate laws of Markov processes on path spaces. The term pinning construction in the title of this section is motivated by the particular example to be studied later, namely by the problem of pinning a Brownian motion on a manifold M down to a submanifold $L \subset M$ by restricting the heat kernel of M to L which gives a proper family of integral operators on L .

The construction of the approximating processes is done in two steps: Given a partition \mathcal{P} of the time interval and a family S of positive integral operators on L we construct in natural way measures $\mathbb{P}_{L,\mathcal{P}}^x$ on the finite product $L^{\mathcal{P}}$. These can be extended to measures $\mathbb{P}_{\mathcal{P}}^x$ on the space of M -valued functions on $[0, 1]$ where $M \supset L$ by allowing the path to make excursions into the surrounding space M in the partition intervals. In this second step there is a choice of various interpolation schemes, ranging from step functions over geodesic interpolation to interpolation via Brownian bridges.

Having established the approximation of the limit semigroup in the functional analytic sense by Chernoff's theorem (which in probabilistic language amounts to convergence of the finite dimensional marginals) the additional problem is to prove tightness of the laws of the approximating processes. This again has two parts. The first part is tightness under stepwise or otherwise trivial interpolation. In the next subsection we improve quite general criteria of (Ethier and Kurtz, 1986) for tightness over the space $D_M[0, 1]$ which extend similar results for $C_{\mathbb{R}^m}[0, 1]$ given by (Stroock and Varadhan, 1979). The second part is to prove tightness over $C_M[0, 1]$ for the interpolation by Brownian bridges which requires additional tools in the case of general manifolds.

Notation. Denote by $|\mathcal{P}|$ the *mesh* of a partition \mathcal{P} , i.e. the length of the longest partition interval. For locally compact separable metric space L let $\widehat{C}(L)$ denote the Banach space of continuous functions vanishing at infinity. Note that a Feller semigroup on $\widehat{C}(L)$ is the transition semigroup of a (strong Markov) process with càdlàg paths, cf. (Ethier and Kurtz, 1986), Theorem 3.2.7, p. 169.

3.1. DISCRETE TIME APPROXIMATIONS OF FELLER PROCESSES

Theorem 2 below shows that if a proper family of integral operators on a locally compact separable metric space is Chernoff equivalent to a Feller semigroup then, for any sequence of partitions whose mesh converges to

0, the associated measures over the piecewise constant paths converge weakly for the Skorokhod topology to the law of the corresponding Feller process. A similar result for uniform partitions and families of Markov operators is Theorem 3.2.6 in (Ethier and Kurtz, 1986).

DEFINITION 3. *Let (L, d) be a metric space and let S a one-parameter family of integral operators*

$$S(t)f(x) = \int \rho(t, x, dy)f(y)$$

where $\rho(t, x, -)$ forms a finite nonnegative Borel measure on L for all $t > 0$, $x \in L$ such that $\rho(t, x, L)$ is bounded in x for each t . We call S a pinning family.

Let $\mathcal{P} := \{0 = t_0 < t_1 \cdots < t_r = 1\}$ be a partition of the interval $[0, 1]$. For every $x \in L$ we define a finite measure on the discrete time 'path space' $L^{\mathcal{P}}$ by

$$\begin{aligned} & \mathbb{P}_{L, \mathcal{P}}^x(A_1 \times \cdots \times A_r) \\ &= \int_{A_1} \cdots \int_{A_r} \rho(t_1 - t_0, x, dy_1) \cdots \rho(t_r - t_{r-1}, y_{r-1}, dy_r). \end{aligned} \tag{7}$$

Also we denote by $\bar{\mathbb{P}}_{L, \mathcal{P}}^x$ the unique measure on the space $D_L[0, 1]$ of càdlàg L -valued paths whose projection to $L^{\mathcal{P}}$ is $\mathbb{P}_{L, \mathcal{P}}^x$ and which is concentrated on the set of paths which are constant on each of the partition intervals $[t_i, t_{i+1})$.

In other words $\bar{\mathbb{P}}_{L, \mathcal{P}}^x$ is the image of $\mathbb{P}_{L, \mathcal{P}}^x$ under the canonical embedding of $L^{\mathcal{P}}$ into $D_L[0, 1]$.

Remarks. 1. The measures $\rho(t, x, -)$ and hence the measures $\mathbb{P}_{L, \mathcal{P}}^x$ and $\bar{\mathbb{P}}_{L, \mathcal{P}}^x$ are only finite but not necessarily probability measures. Nevertheless we shall use the usual topology of weak convergence of measures which is induced by the duality with bounded continuous functions. If the $\rho(t, -, -)$ actually are probability kernels we can interpret $\mathbb{P}_{L, \mathcal{P}}^x$ and $\bar{\mathbb{P}}_{L, \mathcal{P}}^x$ as the laws of two Markov process starting in $x \in L$ with state space L and time parameter set \mathcal{P} and $[0, 1]$, respectively.
2. Up to section 5 the families S will be proper. In section 6 we apply Definition 3 to non proper families.

THEOREM 2. *Let (L, d) be a locally compact and separable metric space. Let $S = (S(t))$ be a pinning family. Suppose that S is proper and Chernoff equivalent on $\hat{C}(L)$ to the Feller semigroup e^{tA} . Then for*

every $x \in L$ and for every sequence of partitions \mathcal{P}_k with $|\mathcal{P}_k| \rightarrow 0$ the associated measures $\bar{\mathbb{P}}_{L,\mathcal{P}_k}^x$ converge weakly over the space $D_L[0,1]$ to the law of the Feller process X starting in x with generator A .

A key observation is the following routine connection between Markov chains and martingales.

LEMMA 3. *Let $Y_i, i = 0, \dots, r$ be a (non homogeneous) Markov chain with transition operators S_i . Then for every real bounded measurable function f the process*

$$f(Y_j) - \sum_{i < j} (S_i - I)f(Y_i) \quad (8)$$

$j = 0, \dots, r$ is a martingale with respect to the natural filtration of (Y_i) .

Proof: (of the theorem) 1. First let us reduce the proof to the normalized case where each $S(t)$ is a Markov transition operators. Since e^{tA} is a Markovian semigroup we have $1 = e^{tA}1$ and $A1 = 0$. Thus by the assumption of Chernoff equivalence and (4) we get

$$\|S(t)1 - 1\| = o(t).$$

The total mass $\rho(t, x, L)$ of the measure $\rho(t, x, -)$ is equal to $S(t)1(x)$. Hence if we consider the probability measure

$$\tilde{\rho}(t, x, -) = \frac{\rho(t, x, -)}{\rho(t, x, L)}$$

the associated Markov operator $\tilde{S}(t)$ differs in operator norm from $S(t)$ only by the order $o(t)$. In particular according to Lemma 1 the families S and \tilde{S} are Chernoff equivalent. From (7) we see that replacing the family S by the family \tilde{S} changes the associated measure $\bar{\mathbb{P}}_{L,\mathcal{P}}^x$ in total variation norm only by $o(|\mathcal{P}_k|)$. This implies a fortiori that the original sequence and the corresponding sequence of probability measures have the same weak limit.

2. Now assume that the measures under consideration are probability measures. For each k we consider the family of operators $T_k(s)$ which is defined by

$$T_k(s) = \frac{S(t_{j+1} - t_j) - I}{t_{j+1} - t_j} \text{ for } s \in [t_j, t_{j+1}). \quad (9)$$

Let $X_{\mathcal{P}_k}^x$ denote the process with law $\bar{\mathbb{P}}_{L,\mathcal{P}_k}^x$. Then

$$\begin{aligned} \sum_{i < j} (S(t_{i+1} - t_i) - I)f(X_{\mathcal{P}_k}^x(t_i)) &= \sum_{i < j} T_k(t_i)f(X_{\mathcal{P}_k}^x(t_i))(t_{i+1} - t_i) \\ &= \int_0^{t_j} T_k(s)f(X_{\mathcal{P}_k}^x(s))ds. \end{aligned} \quad (10)$$

For every function of the form $f = e^{aA}g$ we know from the assumption of Proposition 1 that $T_k(s)f(z) \rightarrow Af(z)$ uniformly in $s \in [0, 1], z \in L$ as $k \rightarrow \infty$. Fix $k \in \mathbb{N}$ and $a > 0$, and let f be of the form $f = e^{aA}g$, $g \in \widehat{C}(L)$. Put

$$M_k^f(t) = f(X_{\mathcal{P}_k}^x(t_j)) - \sum_{i < j} (S(t_{i+1} - t_i) - I)f(X_{\mathcal{P}_k}^x(t_i)) \text{ for } t \in [t_j, t_{j+1})$$

and

$$Z_k^f(t) = M_k^f(t) - f(X_{\mathcal{P}_k}^x(t)). \quad (11)$$

Then by Lemma 3 the process $M_k^f(t)$ is a martingale with respect to the natural filtration (\mathcal{G}_t^k) of $(X_{\mathcal{P}_k}^x(t))$, because $\mathcal{G}_t^k = \mathcal{G}_{t_j}^k$ for $t \in [t_j, t_{j+1})$. Moreover by (10) and the above uniform convergence there is a finite deterministic constant C_f which depends on f such that

$$\sup_k |Z_k^f(t)| = \sup_k \left| \int_0^{t_j} T_k(s)f(X_{\mathcal{P}_k}^x(s))ds \right| \leq C_f t.$$

The set of all functions of the form $f = e^{aA}g$ is uniformly dense in $\widehat{C}(L)$. Thus Theorem 9.4 of chapter 3 of (Ethier and Kurtz, 1986) can be applied to the algebra $C_a = \widehat{C}(L)$ and we conclude that for each $f \in \widehat{C}(L)$ the sequence of processes $f \circ X_{\mathcal{P}_k}^x$ is uniformly tight in $D_{\mathbb{R}}[0, 1]$.

Moreover the finite dimensional marginals of the processes $X_{\mathcal{P}_k}^x$ converge to the corresponding marginals of the process X^x by Proposition 1. Hence Corollary 9.3 of Chapter 3 of (Ethier and Kurtz, 1986) gives the assertion. \blacksquare

In our applications we are interested in weak convergence over the space of continuous rather than càdlàg paths. For the corresponding transfer between these settings the following Lemma is useful:

LEMMA 4. *Let (L, d) be a separable metric space. For a function $\varphi : [0, 1] \rightarrow L$ and $\delta > 0$ let*

$$w(\varphi, \delta) := \sup\{d(\varphi(s), \varphi(t)) : s, t \in [0, 1], |s - t| < \delta\}.$$

Let the sequence (\mathbb{P}_k) of finite measures over $D_L[0, 1]$ converge weakly to a finite measure \mathbb{P}_0 which is concentrated on $C_L[0, 1]$. Then for all $\varepsilon > 0$ one has, as $\delta \downarrow 0$,

$$\limsup_{k \in N} \mathbb{P}_k \{ \omega \in D_L[0, 1] : w(\omega, \delta) > \varepsilon \} \longrightarrow 0. \quad (12)$$

Proof: Since the total mass of the \mathbb{P}_k converges to the total mass of \mathbb{P}_0 we may assume that we deal with probability measures. A sequence in $D_L[0, 1]$ which converges in the topology of this space to an element of $C_L[0, 1]$ actually converges uniformly on $[0, 1]$, cf. (Ethier and Kurtz, 1986), Lemma 3.10.1. By Skorokhod representation (Ethier and Kurtz, 1986), Theorem 3.1.8, there are a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and processes $X^k, k \geq 0$ on this space such that for each k , X^k has law \mathbb{P}_k and $X^k \rightarrow X^0$ a.s. in $D_L[0, 1]$ and hence also a.s. uniformly on $[0, 1]$. Since X^0 has a.s. (uniformly) continuous paths we have for each $\varepsilon > 0$

$$\mathbb{Q}\{w(X^0, \delta) > \varepsilon\} \longrightarrow 0.$$

Because of the a.s. uniform convergence of the paths this implies

$$\limsup_{k \in N} \mathbb{Q}\{w(X^k, \delta) > \varepsilon\} \longrightarrow 0.$$

This is equivalent to (12). ■

3.2. CONTINUOUS INTERPOLATIONS

We assume now that L is embedded into another locally compact space M . We construct a net of measures $\mathbb{P}_{\mathcal{P}}^x$ on the path space $C_M[0, 1]$ associated to the family S . It is indexed by the *finite* partitions \mathcal{P} of $[0, 1]$ and depends on a starting point $x \in L$. The marginals of these measures on $M^{\mathcal{P}}$ are concentrated on $L^{\mathcal{P}}$ and given by the measures $\mathbb{P}_{L, \mathcal{P}}^x$ introduced in Definition 3. In the partition intervals we use an 'interpolation family':

DEFINITION 4. *A family $\mathcal{Q} := \{\mathbb{Q}_{s,t}^{x,y} : x, y \in L, 0 \leq s < t \leq 1\}$ of probability measures on the path space $C_M[s, t]$ such that*

$$\mathbb{Q}_{s,t}^{x,y}(\{\omega : \omega(s) = x, \omega(t) = y\}) = 1$$

is called interpolating family.

We combine the interpolating family $\mathbb{Q}_{t,s}^{x,y}$ with a measure of the form $\mathbb{P}_{L, \mathcal{P}}^x$ to arrive at a path measure as follows: Every $\omega \in C_M[0, 1]$ can be identified with a unique m -tuple

$$\underline{\omega} := (\omega_1, \dots, \omega_r) \in C_M[0, t_1] \times \dots \times C_M[t_{r-1}, 1]$$

which satisfies $\omega_j(t_j) = \omega_{j+1}(t_j)$ for all $j \in \{1, \dots, r-1\}$. Using this identification we define the measure $\mathbb{P}_{\mathcal{P}}^x$ on $C_M[0, 1]$ by

$$\mathbb{P}_{\mathcal{P}}^x(d\omega) = \int_{M^{\mathcal{P}}} \mathbb{P}_{L, \mathcal{P}}^x(dy_1 \times \dots \times dy_r) \mathbb{Q}_{0, t_1}^{x, y_1}(d\omega_1) \dots \mathbb{Q}_{t_{r-1}, 1}^{y_{r-1}, y_r}(d\omega_r). \quad (13)$$

Here $\mathbb{P}_{L,\mathcal{P}}^x$ is considered as a measure on $M^{\mathcal{P}}$. Evaluation at the partition points gives a canonical projection $\pi_{\mathcal{P}}$ of the path space $C_M[0, 1]$ to $M^{\mathcal{P}}$. Then under this projection, the measure $\mathbb{P}_{\mathcal{P}}^x$ has the marginal measure $\mathbb{P}_{L,\mathcal{P}}^x$. In particular both measures have the same total mass.

We give three examples for possible interpolating families.

Examples. (1) L-Geodesic Interpolation: Let L be a Riemannian manifold. The family \mathcal{Q} is given by the point mass $\mathbb{Q}_{s,t}^{x,y} := \delta_{\gamma_{x,y,s,t}^L}$, where

$$\gamma_{x,y,s,t}^L(u) := \gamma_{x,y}^L\left(\frac{u-s}{t-s}d_L(x,y)\right)$$

and $\gamma_{x,y}^L$ is an arbitrary shortest geodesic in L connecting x and y parametrized by arc length. The measures constructed by L -geodesic interpolation are supported by the path space $C_L[0, 1]$.

(2) M-Geodesic Interpolation: Assume that L is isometrically embedded into the manifold M . The family \mathcal{Q} is given by the point mass $\mathbb{Q}_{s,t}^{x,y} := \delta_{\gamma_{x,y,s,t}^M}$, where

$$\gamma_{x,y,s,t}^M(u) := \gamma_{x,y}^M\left(\frac{u-s}{t-s}d_M(x,y)\right)$$

and $\gamma_{x,y}^M$ is an arbitrary shortest geodesic in M connecting x and y parametrized by arc length.

(3) Brownian Bridge Interpolation: Here $\mathbb{Q}_{s,t}^{x,y}$ denotes the measure of a *Brownian bridge* in M starting at time s at x and ending up at time t at y .

For the sake of definiteness we fix now the general setting for the sequel. Nevertheless many of our general arguments could be adapted to other similar situations.

General Assumption. We assume that the Riemannian manifold L is smooth, closed (i.e. compact without boundary) of dimension l and isometrically embedded in to the Riemannian manifold M of dimension m .

Here is the consequence of Theorem 2 in our context. This will be the central general tool in the last two sections.

THEOREM 3. *Under the 'General Assumption' above let S be a proper pinning family which is Chernoff equivalent on the Banach space $C(L)$ to the semigroup (e^{tA}) of a diffusion processes, i.e. a (strong Markov) Feller process with continuous paths on L . Let \mathcal{Q} be either L -geodesic, M -geodesic or Brownian bridge interpolation or another interpolating family for which the assumptions of Lemma 5 below hold. Then for every $x \in L$ and for every sequence of partitions \mathcal{P}_k with $|\mathcal{P}_k| \rightarrow 0$*

the associated measures $\mathbb{P}_{\mathcal{P}_k}^x$ obtained by (13) converge weakly over the space $C_M[0, 1]$ to the law of the process X starting in x with generator A .

Proof: (1) We begin with M -geodesic interpolation. Fix $\delta > 0$. Choose k large enough such that $|\mathcal{P}_k| < \delta$. Let $s < t$ with $t - s < \delta$ and a path ω be given which is geodesic in the intervals of the partition $\mathcal{P}_k = \{t_0 < \dots < t_{r_k}\}$. Choose the indices l, u such that $t_l := \max\{\tau \in \mathcal{P}_k : \tau \leq s\}$ and $t_u := \min\{\tau \in \mathcal{P}_k : t \leq \tau\}$. By the construction of geodesic interpolation, we have

$$\begin{aligned} & d_M(\omega(s), \omega(t)) \\ & \leq d_M(\omega(s), \omega(t_{l+1})) + d_M(\omega(t_{l+1}), \omega(t_{u-1})) + d_M(\omega(t_{u-1}), \omega(t)) \\ & \leq d_M(\omega(t_{l-1}), \omega(t_l)) + d_M(\omega(t_{l+1}), \omega(t_{u-1})) + d_M(\omega(t_{u-1}), \omega(t_u)) \\ & \leq 3 \max\{d_M(\omega(t_i), \omega(t_j)) : t_i, t_j \in \mathcal{P}_k, |t_j - t_i| < \delta\}. \end{aligned} \tag{14}$$

The measures $\bar{\mathbb{P}}_{L, \mathcal{P}_k}^x$ of Definition 3 and $\mathbb{P}_{\mathcal{P}_k}^x$ defined in (13) have the same marginal measure $\mathbb{P}_{L, \mathcal{P}_k}^x$ on $M^{\mathcal{P}_k}$ resp $L^{\mathcal{P}_k}$. Let us denote by w_M and w_L respectively the modulus of continuity introduced in Lemma 4 with respect to the metrics d_M resp d_L . Then

$$\begin{aligned} & \mathbb{P}_{\mathcal{P}_k}^x \{ \omega \in C_M[0, 1] : w_M(\omega, \delta) > \varepsilon \} \\ & \leq \mathbb{P}_{L, \mathcal{P}_k}^x \{ \omega \in L^{\mathcal{P}_k} : \max\{d_M(\omega(t_i), \omega(t_j)) : |t_j - t_i| < \delta\} > \varepsilon/3 \} \\ & \leq \mathbb{P}_{L, \mathcal{P}_k}^x \{ \omega \in L^{\mathcal{P}_k} : \max\{d_L(\omega(t_i), \omega(t_j)) : |t_j - t_i| < \delta\} > \varepsilon/3 \} \\ & \leq \bar{\mathbb{P}}_{L, \mathcal{P}_k}^x \{ \omega \in D_L[0, 1] : w_L(\omega, \delta) > \varepsilon/3 \}. \end{aligned}$$

According to Theorem 2 the sequence $(\bar{\mathbb{P}}_{L, \mathcal{P}_k}^x)_k$ converges weakly over the space $D_L[0, 1]$ to the law of the continuous process X . Thus by Lemma 4 we get for each $\varepsilon > 0$

$$\lim_{\delta \rightarrow 0} \limsup_{k \rightarrow \infty} \mathbb{P}_{\mathcal{P}_k}^x \{ \omega \in C_M[0, 1] : w_M(\omega, \delta) > \varepsilon \} = 0. \tag{15}$$

This implies tightness of the sequence $(\mathbb{P}_{\mathcal{P}_k}^x)_k$ over the space $C_M[0, 1]$ and clearly the law of X is the only possible limit point.

- (2) The proof for L -geodesic interpolation is completely analogous taking d_L, w_L instead of d_M, w_M .
- (3) For the Brownian bridge interpolation the result now follows from Lemma 5 and Proposition 2 below. ■

We prove that Brownian bridge interpolation leads to the same limit measure as geodesic interpolation. If $M = \mathbb{R}^m$ a much easier proof of this fact was given in (Smolyanov, v. Weizsäcker and Wittich, 2000).

LEMMA 5. Let \mathcal{Q} be an interpolating family such that for $\varepsilon > 0$ small enough, we have for all $\alpha > 0$ an increasing function $g_\alpha : [0, \alpha] \rightarrow \mathbb{R}_0^+$ with $\lim_{u \rightarrow 0} g_\alpha(u) = 0$ such that

$$\mathbb{Q}_{s,t}^{x,y}(\Gamma_{x,y,s,t}^M(\alpha)) \leq (t-s)g_\alpha(t-s)$$

uniformly in

$$R_\varepsilon := \{(x, y) \in M \times M : d_M(x, y) < \varepsilon\}$$

where

$$\Gamma_{x,y,s,t}^M(\alpha) := \{\omega \in C_M[s, t] : \sup_{s \leq u \leq t} d_M(\omega(u), \gamma_{x,y,s,t}^M(u)) \geq \alpha\}.$$

Then the sequence $\mathbb{W}_{\mathcal{P}_k}^x$ constructed from S and \mathcal{Q} converges to the same limit measure as the sequence constructed from S by M -geodesic interpolation.

Proof: Let ν_k and ρ_k be measures on $C_M[0, 1]$ constructed from \mathcal{Q} and M -geodesic interpolation respectively. For each path $\omega \in C_M[0, 1]$ let $\varphi_k(\omega)$ be a M -geodesic interpolation of the restriction $\omega|_{\mathcal{P}_k}$. Then we have the relation $\rho_k = \nu_k \circ \varphi_k^{-1}$. Moreover

$$\begin{aligned} & \nu_k \{ \omega : \sup_{0 \leq u \leq 1} d_M(\omega(u), \varphi_k(\omega)(u)) > \alpha \} \\ & \leq \sum_{l=1}^{r_k} \nu_k \{ \omega : \sup_{t_{l-1} \leq u \leq t_l} d_M(\omega(u), \varphi_k(\omega)(u)) > \alpha \} \\ & \leq \sum_{l=1}^{r_k} \int \mathbb{P}_{L, \mathcal{P}_k}^x(dy) \mathbb{Q}_{t_{l-1}, t_l}^{y_{l-1}, y_l} \{ \Gamma_{y_{l-1}, y_l, t_{l-1}, t_l}^M(\alpha) \} \\ & \leq \mathbb{P}_{L, \mathcal{P}_k}^x \{ w(\omega, |\mathcal{P}_k|) > \varepsilon \} + \sum_{l=1}^{r_k} \sup_{d_M(y_{l-1}, y_l) \leq \varepsilon} \mathbb{Q}_{t_{l-1}, t_l}^{y_{l-1}, y_l} \{ \Gamma_{y_{l-1}, y_l, t_{l-1}, t_l}^M(\alpha) \} \\ & \leq \mathbb{P}_{L, \mathcal{P}_k}^x \{ w(\omega, |\mathcal{P}_k|) > \varepsilon \} + \sum_{l=1}^{r_k} (t_l - t_{l-1}) g_\alpha(t_l - t_{l-1}). \end{aligned}$$

Here both terms are arbitrarily small for large k and small α : The first by Lemma (4) and the second due to our assumption. Hence for every uniformly continuous function $f : C_M[0, 1] \rightarrow \mathbb{R}$ we get

$$\lim_{k \rightarrow \infty} \left| \int f(\omega) d\nu_k - \int f(\omega) d\rho_k \right| \leq \int |f(\omega) - f(\varphi_k(\omega))| d\nu_k = 0$$

which implies that the sequences (ν_k) and (ρ_k) have the same limit points in law. \blacksquare

Remark. 1. Instead of M -geodesic interpolation we could have used in Lemma 5 any other interpolating family as reference for which the pinning measures are known to be tight. In the proof of Theorem 3 the geodesic interpolation was used to get the second inequality in (14). Clearly a uniform estimate of the form

$$d_M(\omega(s), \omega(t)) \leq h(d_M(\omega(t_i), \omega(t_{i+1}))) \text{ for } s, t \in (t_i, t_{i+1})$$

where $\lim_{\varepsilon \rightarrow 0} h(\varepsilon) = 0$ would have been sufficient. It is easy to construct other interpolation schemes on more general metric spaces which satisfy such a condition.

2. By LeCam's Theorem, see (Dudley, 1989), 11.5.3 Theorem, p. 316, convergence of the respective sequence implies its uniform tightness.

Finally, we use a uniform Large-Deviation result about Brownian bridges (Wittich, 2005) to conclude the corresponding part of Theorem 3 from Lemma 5.

PROPOSITION 2. *The Brownian bridge interpolation family \mathcal{Q} on M satisfies the assumption of Lemma 5 with $g_\alpha(u) := 2e^{-\chi\alpha^2/u}/u$ for some $\chi > 0$.*

Proof: As in the proof at the end of the previous section, let $r_M := \inf_{x \in L} r_M(x)$ where $r_M(x) > 0$ is the largest number such that the geodesic balls $B(x, r)$ are *strongly convex* in the sense of (do Carmo, 1992), 3.4, p. 74 for all $r < r_M(x)$. Let $\varepsilon < r_M/2$. By the Large Deviation result from (Wittich, 2005), there are $\chi > 0$ and $\alpha_0 > 0$ such that for all $0 < \alpha < \alpha_0$ we have

$$\mathbb{Q}_{s,t}^{x,y}(\Gamma_{x,y,s,t}^M(\alpha)) \leq 2 \exp\left(-\frac{\chi\alpha^2}{t-s}\right)$$

as $t - s \rightarrow 0$, uniformly in R_ε . This completes the proof. ■

4. Gaussian Integrals

In the sequel, proper families as described in section 2 will be constructed by families of integral operators. In order to compute the derivative at zero – and therefore the Chernoff equivalence class – of such a family we first review some facts about the short time asymptotic of Gaussian integrals from (Smolyanov, v. Weizsäcker and Wittich,

2003). We introduce a degree d on the space of space-time polynomials such that the short-time contribution of a monomial p either vanishes or is of order $t^{d(p)}$. Using this notion, we reformulate Wick's formula in an algebraic way and conclude Corollary 1. It states that in our situation the only relevant terms are of homogeneous degree one. For the proofs of the results in this section cf. (Smolyanov, v. Weizsäcker and Wittich, 2003), p. 351-354.

4.1. WICK'S FORMULA

Let $t > 0$. By Fubini's theorem and using the fact that the Gaussian integral solves the heat equation, we obtain the following result also known as *Wick's formula*.

LEMMA 6. *For $k \in \mathbb{Z} \times \mathbb{N}_0^n$ define*

$$d(k) := k_0 + \frac{1}{2}(k_1 + \dots + k_n). \quad (16)$$

Let

$$p_k(t, \xi) = t^{k_0} \xi_1^{k_1} \dots \xi_n^{k_n}$$

be a monomial in (t, ξ) such that $d(k) \geq 0$. Then for

$$\mathcal{G}_t(p_k) := \int_{\mathbb{R}^n} \frac{e^{-\frac{|\xi|^2}{2t}}}{\sqrt{2\pi t}^n} p_k(t, \xi) \, d\xi \quad (17)$$

we obtain

$$\mathcal{G}_t(p_k) = \begin{cases} 0 & (k_1, \dots, k_n) \notin (2\mathbb{N}_0)^n \\ t^{d(k)} \prod_{i=1}^n (k_i - 1)!! & \text{else} \end{cases}.$$

Here we use the standard notation $(2n - 1)!! = 1 \cdot 3 \cdot \dots \cdot (2n - 1)$.

Let now \mathcal{L} denote the real algebra generated by all monomials of the form $p_k(t, x)$, $k \in \mathbb{Z} \times \mathbb{N}^n$. The map

$$D : \mathcal{L} \rightarrow \frac{1}{2}\mathbb{Z}$$

given by $D(p_k) := d(k)$ induces a *grading* $\mathcal{L} = \bigoplus_{s \in \frac{1}{2}\mathbb{Z}} \mathcal{L}^s$, where $\mathcal{L}^s := \langle p_k : d(k) = s \rangle$ denotes the subspace of homogeneous elements of degree s . We consider the *associated filtration* by ideals $\mathcal{I}^r := \bigoplus_{s \geq r} \mathcal{L}^s$. Let furthermore $f \mapsto [f]$ denote the quotient map

$$q : \mathcal{I}^0 \rightarrow \mathcal{I}^0 / \mathcal{I}^{3/2}$$

and Q the projection onto the subalgebra generated by monomials p_k with $(k_1, \dots, k_n) \in (2\mathbb{N}_0)^n$.

If in particular p_k is a monomial with $k_0 = 0$ then $Q[p_k] = 0$ unless $p_k \in \{1, x_1^2, \dots, x_n^2\}$. The following Proposition about the short time asymptotic of Gaussian integrals is a simple consequence of Lemma 6.

PROPOSITION 3. *Let $f \in \mathcal{I}^0$. Then $\lim_{t \rightarrow 0} \mathcal{G}_t(f)$ exists and we have asymptotically*

$$\mathcal{G}_t(f) = \mathcal{G}_t(Q[f]) + o(t) \quad (18)$$

as $t \rightarrow 0$. More explicitly, if $f = f_0 + f_{1/2} + f_1 + \dots$ is the decomposition of f into homogeneous elements, then

$$\mathcal{G}_t(f) = \mathcal{G}_t(Qf_0) + \mathcal{G}_t(Qf_1) + o(t). \quad (19)$$

Remark. q is a ring homomorphism, whereas Q is not.

In the sequel we will use the following fact concerning quotients of Gaussian integrals.

COROLLARY 1. *Let $f, h \in \mathcal{L}$ have the homogeneous components*

- (i) $f = f_0 + f_{1/2} + f_1$ and $f_0 \in \mathbb{R}$ is constant.
- (ii) $h = 1 + h_1$.

Then, as $t \downarrow 0$,

$$\frac{\mathcal{G}_t(fh)}{\mathcal{G}_t(h)} - \frac{\mathcal{G}_0(fh)}{\mathcal{G}_0(h)} = \mathcal{G}_t(f_1) + o(t). \quad (20)$$

In order to make use of the above discussion we have to show how it applies to more general situations. We observe first that the polynomial short time asymptotic of a Gaussian integral is in some sense independent of the domain of integration. From this we draw the following conclusion for the polynomial short time asymptotic of more general functions (see (Smolyanov, v. Weizsäcker and Wittich, 2003), Corollary 3):

COROLLARY 2. *Let $U \subset \mathbb{R}^n$ be a bounded open neighbourhood of the origin. For $r \in \mathbb{N}$ and $f \in C^r(U) \cap C(\overline{U})$ denote the Taylor polynomial of f up to order $r-1$ (the $(r-1)$ -jet) around 0 by \hat{f} . For every $k \in \mathbb{Z}$ such that $t^k \hat{f} \in \mathcal{I}^0$ we have asymptotically as $t \rightarrow 0$*

$$\int_U \frac{e^{-\frac{|\xi|^2}{2t}}}{\sqrt{2\pi t}^n} t^k f(\xi) d\xi = \mathcal{G}_t(t^k \hat{f}) + O(t^{k+\frac{r}{2}})$$

where the constant in the error term depends on f only via the maximal Taylor coefficient of f of order r in a small neighbourhood of 0.

Remark. In the subsequent asymptotic computations of integrals, we will denote a C^r -function and its Taylor expansion by the same symbol omitting the hat introduced above. For example, for a function $\hat{f} \in C^r(L)$ such that $t^k \hat{f} \in \mathcal{I}^0$, $[t^k \hat{f}]$ denotes the equivalence class of $t^k \hat{f}$ in $\mathcal{I}^0/\mathcal{I}^{3/2}$.

5. Proper Families Equivalent to the Heat Semigroup on Manifolds

5.1. INTRODUCTION

Again let L be a closed smooth manifold and vol_L the corresponding Riemannian volume measure. Our examples in this section follow the same pattern: Consider a family of smooth integral kernels $q_t(x, y) \in C_+^\infty(L \times L)$, $t > 0$ and the associated operators

$$S(t)f(x) = \int_L q_t(x, y)f(y)\text{vol}_L(dy). \quad (21)$$

on the Banach space $C(L)$. We introduce the corresponding (normalized) Markov operators

$$T(t)f(x) = \frac{\int_L q_t(x, y)f(y)\text{vol}_L(dy)}{\int_L q_t(x, z)\text{vol}_L(dz)}. \quad (22)$$

We also assume resp. verify that there is a function $D \in C(L)$ such that the denominator in (22) satisfies

$$b(t, x) := \int_L q_t(x, y)\text{vol}_L(dy) = e^{tD(x)} + o(t) \quad (23)$$

uniformly in $x \in L$ as $t \downarrow 0$.

PROPOSITION 4. *Under the assumption (23) the family $(T(t))$ given by (22) is proper if and only if the operator family $(B(t))$ defined by*

$$B(t)f(x) = \int_L e^{-tD(x)} q_t(x, y)f(y)\text{vol}_L(dy) \quad (24)$$

is proper and in this case they are Chernoff equivalent.

Proof: We have $B(t)f(x) = c(t)T(t)f(x)$ where $c(t)$ is the operator of multiplication with the function

$$x \mapsto b(t, x)e^{-tD(x)}$$

By (23) and Lemma 7 below we can apply Lemma 2 to the operators $c(t)$ and get the assertion. \blacksquare

Actually the corresponding semigroup will be always the heat semigroup $(e^{\frac{t}{2}\Delta_L})$.

LEMMA 7. *Let $k \in C(L \times L)$, $k \geq 0$ and consider the integral operator*

$$I_k f(x) := \int_L k(x, y) f(y) \text{vol}_L(dy).$$

Then $I_k : C(L) \rightarrow C(L)$ is a bounded operator with norm

$$\|I_k\| = \sup_{x \in L} \left| \int_L k(x, y) \text{vol}_L(dy) \right| = \|I_k 1\|_\infty.$$

Therefore we can infer from Theorem 3 convergence of the measures constructed by the pinning construction starting with the either of the families $(T(t))$ or $(B(t))$ to the law of the diffusion process generated by A .

5.2. FIRST EXAMPLES

In this subsection we follow, with a couple of minor corrections, the exposition in (Smolyanov, v. Weizsäcker and Wittich, 2003) in order to motivate the subsequent calculations in subsections 5.3 - 5.5. We fix the Laplace-Beltrami-operator to be non-positive as our choice of sign and consider the heat semigroup $(e^{t\Delta_L/2})$ on the Banach space $C(L)$.

Let $d_L(-, -)$ denote the distance function on L . Consider the pseudo-Gaussian kernel

$$k_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{d_L(x, y)^2}{2t}}.$$

The first step is to consider the family of associated (normalized) Markov-operators

$$T(t)f(x) := \frac{\int_L e^{-\frac{d_L(x, y)^2}{2t}} f(y) \text{vol}_L(dy)}{\int_L e^{-\frac{d_L(x, y)^2}{2t}} \text{vol}_L(dy)}. \quad (25)$$

By the smoothness of the heat kernel on L , the subspace $C^3(L) \subset C(L)$ contains the image $e^{\frac{a}{2}\Delta_L} C(L)$ for each $a > 0$. Thus we prove Chernoff equivalence of the family above and the heat semigroup if we can show that

$$\lim_{t \rightarrow 0} \frac{T(t)f - f}{t} = \frac{1}{2} \Delta_L f \quad (26)$$

for all $f \in C^3(L)$. As noted above, we may restrict ourselves to integration over an arbitrary open neighbourhood $U(x)$ instead of over all of L . We choose $U(x)$ so small that we can use the exponential map to construct a *normal coordinate system*

$$\exp_x^L : V(0) \rightarrow U(x).$$

Let $\xi := \exp_x^L(y)$. Then we have

$$d_L(x, y) = |\xi|$$

in these coordinates. Furthermore

$$\text{vol}_L(dy) = \sqrt{\det g}(\xi) d\xi$$

where g is the metric tensor. Therefore we obtain

$$(T(t)f - f)(x) = \frac{\mathcal{G}_t(\sqrt{\det g} f)}{\mathcal{G}_t(\sqrt{\det g})}(0) - \frac{\mathcal{G}_0(\sqrt{\det g} f)}{\mathcal{G}_0(\sqrt{\det g})}(0).$$

But now $\sqrt{\det g}$ is infinitely differentiable due to our assumptions on the manifold and f is in $C^3(L)$. We may thus apply Corollary 2 with $k = 0$ and in the sequel we just have to consider the Taylor expansion of these functions up to second order. The expansion of the metric tensor g we quote from (Roe, 1988), (1.14) Proposition, p. 8

LEMMA 8. *In normal coordinates the Taylor expansion of g is given by*

$$g_{ab}(\xi) = \delta_{ab} + \frac{1}{3} R_{auvb}^L(0) \xi^u \xi^v + O(|\xi|^3),$$

where R^L denotes the curvature tensor of L .

Thus we get for the volume form, cf. e.g. (Smolyanov, v. Weizsäcker and Wittich, 2003), Cor. 4:

COROLLARY 3. *In normal coordinates the Taylor expansion of $\sqrt{\det g}$ is given by*

$$\sqrt{\det g}(\xi) = 1 + \frac{1}{6} R_{auva}^L(0) \xi^u \xi^v + O(|\xi|^3),$$

where R^L denotes the curvature tensor of L .

Therefore the 2-jet of $\sqrt{\det g}$ has exactly the properties required for the function h in Corollary 1. So we apply Corollary 2 and Corollary 1 to the 2-jets of $f \circ \exp_x^L$ and $\sqrt{\det g}$ to get

$$\begin{aligned} (T(t)f - f)(x) &= \mathcal{G}_t(f_1)(0) + O(t^{3/2}) \\ &= \frac{t}{2} \Delta f(0) + O(t^{3/2}). \end{aligned}$$

But in normal coordinates, the Laplacian on L coincides with Δ , since we were assuming our Laplace-Beltrami operator always to be non-positive. Therefore we may write invariantly

$$(T(t)f - f)(x) = \frac{t}{2} \Delta_L f(x) + O(t^{3/2}),$$

for all functions $f \in C^3(L)$. This is a pointwise statement. But inspecting Corollary 2 above shows that due to compactness the remainder is $O(t^{3/2})$ uniformly on L . This finally implies

PROPOSITION 5. *The family $(T(t))$ defined in (25) is Chernoff equivalent to the heat semigroup on L .*

Next we want to omit the denominator in (25) and compensate it by a suitable modification of the kernel. For this we consider the short time asymptotic of this denominator

$$b(t, x) := \int_L e^{-\frac{d_L(x,y)^2}{2t}} \text{vol}_L(dy).$$

LEMMA 9. *Let Scal_L be the scalar curvature of L . Then, uniformly in $x \in L$,*

$$b(t, x) = \sqrt{2\pi t}^l \left(e^{-\frac{t\text{Scal}_L(x)}{6}} + O(t^{3/2}) \right).$$

Proof: Using again the Taylor expansion of the volume form, we get

$$\begin{aligned} b(t, x) &= \sqrt{2\pi t}^l \left(\mathcal{G}_t([\sqrt{\det g}]) + O(t^{3/2}) \right) \\ &= \sqrt{2\pi t}^l \left(1 + \frac{t}{2} \Delta \left(\frac{1}{6} R_{auva}^L \xi^u \xi^v \right)_{\xi=0} + O(t^{3/2}) \right) \\ &= \sqrt{2\pi t}^l \left(1 + \frac{t}{6} R_{auua}^L(0) + O(t^{3/2}) \right) \\ &= \sqrt{2\pi t}^l \left(1 - \frac{t}{6} \text{Scal}_L(x) + O(t^{3/2}) \right). \end{aligned}$$

■

Propositions 4 and 5 now imply

COROLLARY 4. *The family of bounded operators defined by*

$$B(t)f(x) := \frac{1}{\sqrt{2\pi t}^l} \int_L e^{-\frac{d_L(x,y)^2}{2t} + \frac{t\text{Scal}_L(x)}{6}} f(y) \text{vol}_L(dy). \quad (27)$$

is also Chernoff equivalent to the heat semigroup on L .

Finally, it should be noted that there is also a symmetric version of the approximating kernel. By the very same arguments as in Lemma 9 with

$$\hat{b}(t, x) := \int_L e^{-\frac{d_L(x, y)^2}{2t} + \frac{t(\text{Scal}_L(x) + \text{Scal}_L(y))}{12}} \text{vol}_L(dy)$$

instead of $b(t, x)$ we obtain as well:

COROLLARY 5. *The operator family defined by*

$$\hat{B}(t)f(x) := \frac{1}{\sqrt{2\pi t}^l} \int_L e^{-\frac{d_L(x, y)^2}{2t} + \frac{t\text{Scal}_L(x) + t\text{Scal}_L(y)}{12}} f(y) \text{vol}_L(dy)$$

is proper and Chernoff equivalent to the heat semigroup on L .

5.3. THE HEAT EQUATION ON A SUBMANIFOLD VIA THE RESTRICTION OF A PSEUDO-GAUSSIAN KERNEL

We now consider the restriction to L of a pseudo-Gaussian kernel on M , and prove Chernoff equivalence to the heat semigroup on L . This result was communicated to us with a different proof already in (Tokarev, 2001). We will state it in the spirit of the preceding sections as follows:

THEOREM 4. *Let $\phi : L \subset M$ be an isometric embedding of the closed and connected smooth Riemannian manifold L into the smooth Riemannian manifold M . Let $\dim(L) = l$, $\dim(M) = m$. Let Scal_L be the scalar curvature of L . Let τ_ϕ denote the tension vectorfield of the embedding and*

$$\overline{R}_{M/L} := \sum_{a,b=1}^l \langle R^M(e_a, e_b)e_b, e_a \rangle \quad (28)$$

the partial trace of the curvature tensor of M over an arbitrary orthonormal base of $\phi_ TL$. Then the family $(B(t))$ defined by*

$$B(t)f(x) := \frac{e^{t(\frac{\text{Scal}_L}{4} - \frac{|\tau_\phi|^2}{8} - \frac{\overline{R}_{M/L}}{12})(x)}}{\sqrt{2\pi t}^l} \int_L e^{-\frac{d_M(\phi(x), \phi(y))^2}{2t}} f(y) \text{vol}_L(dy) \quad (29)$$

is proper and Chernoff equivalent to the heat semigroup on L .

Specializing this result to embeddings into euclidean space we obtain the following statement.

COROLLARY 6. *If, in Theorem 4, $M = \mathbb{R}^m$ then the family*

$$B_t f(x) := \frac{e^{t(\frac{\text{Scal}_L}{4} - \frac{|\tau_\phi|^2}{8})(x)}}{\sqrt{2\pi t}} \int_L e^{-\frac{|\phi(y) - \phi(x)|^2}{2t}} f(y) \text{vol}_L(dy)$$

is proper and Chernoff equivalent to the heat semigroup on L .

5.4. PROOF OF THEOREM 4

Let $x, y \in L$ and $f \in C^3(L)$. We now want to determine an asymptotic expression for

$$S(t)f(x) := \frac{1}{\sqrt{2\pi t}} \int_L e^{-\frac{d_M(\phi(y), \phi(x))^2}{2t}} f(y) \text{vol}_L(dy). \quad (30)$$

By the arguments above we can reduce the problem to purely local considerations on sufficiently small neighbourhoods $U_L(x)$ and $U_M(\phi(x))$ which are chosen such that $\phi(L) \cap U_M(\phi(x)) = \phi(U_L(x))$. To do so, we consider local normal coordinates

$$\begin{aligned} \exp_x^L : V(0) &\rightarrow U_L(x) \\ \exp_{\phi(x)}^M : W(0) &\rightarrow U_M(x). \end{aligned}$$

Local coordinates for L and M are denoted by $\xi = (\xi^1, \dots, \xi^l)$ and $\eta = (\eta^1, \dots, \eta^m)$ respectively. The local coordinate representation

$$(\exp_{\phi(x)}^M)^{-1} \circ \phi \circ \exp_x^L : V(0) \rightarrow W(0)$$

of ϕ will be denoted by the same letter, i.e.

$$\eta = \phi(\xi) = (\phi^1(\xi), \dots, \phi^m(\xi)).$$

In these local coordinates we obtain

$$S(t)f(x) = \frac{1}{\sqrt{2\pi t}} \int_{V(0)} e^{-\frac{|\phi(\xi)|^2}{2t}} f(\xi) \sqrt{\det g^L}(\xi) d\xi + O(t^{3/2}). \quad (31)$$

We want to apply Proposition 3. To do so we have to make sure that

$$[h(\xi)] = \left[e^{-\frac{|\phi(\xi)|^2 - |\xi|^2}{2t}} \sqrt{\det g^L} f(\xi) \right]$$

really satisfies the assumptions made there. To see this we denote by

$$H_\phi(-, -) = \nabla^L d\phi(-, -), \quad (32)$$

the *Hessian* of the map ϕ , which coincides with the *second fundamental form* of the embedding. Now the necessary input from differential geometry can be summarized in the following proposition.

PROPOSITION 6. Consider points $x, y \in L \subset M$, where L is considered as isometrically embedded by the map ϕ as described above. Let $p \in U(x)$ and $U(x) \subset L$ so small that y and x are joined by a unique minimizing geodesic γ_{xy}^L starting at x . We assume γ_{xy}^L to be parametrized by arc-length. Then

$$\lim_{d_L(x,y) \rightarrow 0} \frac{d_M(x,y)^2 - d_L(x,y)^2}{d_L(x,y)^4} = -\frac{1}{12} \left\| H_\phi(\dot{\gamma}_{xy}^L(0), \dot{\gamma}_{xy}^L(0)) \right\|^2 \quad (33)$$

Proof: In those local coordinates defined above, we have using the Taylor expansion of ϕ around $\xi = 0$:

$$\begin{aligned} \frac{d_M(x,y)^2 - d_L(x,y)^2}{d_L(x,y)^4} &= \frac{|\phi(\xi)|^2 - |\xi|^2}{|\xi|^4} \\ &= \frac{\left| \frac{\partial \phi^\alpha}{\partial \xi^a}(0) \xi^a + \frac{1}{2} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b}(0) \xi^a \xi^b + \frac{1}{6} \frac{\partial^3 \phi^\alpha}{\partial \xi^a \partial \xi^b \partial \xi^u}(0) \xi^a \xi^b \xi^u \right|^2 - |\xi|^2}{|\xi|^4} + O(|\xi|). \end{aligned}$$

Denote the metric on L by g^L and the metric on M by g^M .¹ The fact that ϕ is an isometric embedding is equivalent to the local equation (see (Jost, 1998), p. 29 f.)

$$g_{ab}^L(\xi) = g_{\alpha\beta}^M(\eta) \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b}. \quad (34)$$

Since the metric tensor at the origin of a normal coordinate system is the flat one, we obtain using (34)

$$g_{\alpha\beta}^M \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b}(0) \xi^a \xi^b = g_{ab}^L(0) \xi^a \xi^b = |\xi|^2.$$

Partial differentiation of (34) yields

$$\frac{\partial g_{ab}^L}{\partial \xi^u} = \frac{\partial g_{\alpha\beta}^M}{\partial \eta^\rho} \frac{\partial \phi^\rho}{\partial \xi^u} \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b} + g_{\alpha\beta}^M \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^u} \frac{\partial \phi^\beta}{\partial \xi^b} + \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial^2 \phi^\beta}{\partial \xi^b \partial \xi^u} \right).$$

But at the origin of a normal coordinate system the partial derivative of the metric tensor coincides with its covariant derivative and therefore vanishes. This implies

$$\delta_{\alpha\beta} \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^u} \frac{\partial \phi^\beta}{\partial \xi^b} + \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial^2 \phi^\beta}{\partial \xi^b \partial \xi^u} \right)(0) = 0.$$

¹ Throughout we use freely the Einstein summation convention. From now on greek indices will be used for the summation from 1 to m and latin indices for summation from 1 to l

This means

$$\begin{aligned}
& \frac{d_M(x, y)^2 - d_L(x, y)^2}{d_L(x, y)^4} \\
&= \frac{\delta_{\alpha\beta} \left(\frac{1}{4} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v} + \frac{1}{3} \frac{\partial^3 \phi^\alpha}{\partial \xi^a \partial \xi^b \partial \xi^u} \frac{\partial \phi^\beta}{\partial \xi^v} \right) (0) \xi^a \xi^b \xi^u \xi^v}{|\xi|^4} + O(|\xi|) \\
&= -\frac{1}{12} \delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v} (0) \frac{\xi^a \xi^b \xi^u \xi^v}{|\xi|^4} \\
&\quad + \frac{1}{3} \delta_{\alpha\beta} \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v} + \frac{\partial^3 \phi^\alpha}{\partial \xi^a \partial \xi^b \partial \xi^u} \frac{\partial \phi^\beta}{\partial \xi^v} \right) (0) \frac{\xi^a \xi^b \xi^u \xi^v}{|\xi|^4} + O(|\xi|).
\end{aligned}$$

If we now differentiate (34) twice we obtain at the origin

$$\begin{aligned}
\frac{\partial^2 g_{ab}^L}{\partial \xi^u \partial \xi^v} (0) &= \frac{\partial^2 g_{\alpha\beta}^M}{\partial \eta^\rho \partial \eta^\mu} \frac{\partial \phi^\rho}{\partial \xi^u} \frac{\partial \phi^\mu}{\partial \xi^v} \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b} (0) + \delta_{\alpha\beta} \left(\frac{\partial^3 \phi^\alpha}{\partial \xi^a \partial \xi^u \partial \xi^v} \frac{\partial \phi^\alpha}{\partial \xi^b} \right. \\
&\quad \left. + \frac{\partial^3 \phi^\alpha}{\partial \xi^b \partial \xi^u \partial \xi^v} \frac{\partial \phi^\alpha}{\partial \xi^a} + \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^u} \frac{\partial^2 \phi^\alpha}{\partial \xi^b \partial \xi^v} + \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^v} \frac{\partial^2 \phi^\alpha}{\partial \xi^b \partial \xi^u} \right) (0).
\end{aligned}$$

Using this and relating the partial derivatives of the metric tensor to curvature by using the fact (see Lemma 8) that in normal coordinates the Taylor expansion of g is given by

$$g_{ab}(\xi) = \delta_{ab} + \frac{1}{3} R_{auvb}(0) \xi^u \xi^v + O(|\xi|^3), \quad (35)$$

where R denotes the curvature tensor, we obtain

$$\begin{aligned}
& 2\delta_{\alpha\beta} \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v} + \frac{\partial^3 \phi^\alpha}{\partial \xi^a \partial \xi^b \partial \xi^u} \frac{\partial \phi^\beta}{\partial \xi^v} \right) (0) \xi^a \xi^b \xi^u \xi^v \\
&= \left(\frac{\partial^2 g_{ab}^L}{\partial \xi^u \partial \xi^v} - \frac{\partial^2 g_{\alpha\beta}^M}{\partial \eta^\rho \partial \eta^\mu} \frac{\partial \phi^\rho}{\partial \xi^u} \frac{\partial \phi^\mu}{\partial \xi^v} \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b} \right) (0) \xi^a \xi^b \xi^u \xi^v \\
&= 2 \left(R_{auvb}^L - R_{\alpha\rho\mu\beta}^M \frac{\partial \phi^\rho}{\partial \xi^u} \frac{\partial \phi^\mu}{\partial \xi^v} \frac{\partial \phi^\alpha}{\partial \xi^a} \frac{\partial \phi^\beta}{\partial \xi^b} \right) (0) \xi^a \xi^b \xi^u \xi^v = 0
\end{aligned}$$

due to the symmetries of the curvature tensor (see (Jost, 1998), (3.3.7), p. 129). On the other hand, the remaining term is indeed the Hessian at the origin of a local normal coordinate system (see (Jost, 1998), (3.3.47), p. 138) and since $\xi = (\exp_x^L)^{-1}(y)$ and therefore $\xi/|\xi| = \dot{\gamma}_{xy}^L(0)$ we finally obtain our statement. \blacksquare

We now obtain the following result, first derived with a different proof in (Tokarev, 2001).

PROPOSITION 7. *With the notations of Theorem 4 we have for $f \in C^3(L)$*

$$S(t)f(x) = e^{tD(x)}f(x) + \frac{t}{2}\Delta_L f(x) + O(t^{3/2}), \quad (36)$$

in particular

$$\frac{1}{\sqrt{2\pi t}^l} \int_L e^{-\frac{d_M(\phi(y), \phi(x))^2}{2t}} \text{vol}_L(dy) = e^{tD(x)} + O(t^{3/2}) \quad (37)$$

where

$$D(x) = - \left(\frac{\text{Scal}_L}{4} + \frac{|\tau_\phi|^2}{8} + \frac{\bar{R}_{M/L}}{12} \right) (x).$$

Proof: By (31) and Proposition 6

$$\begin{aligned} S(t)f(x) &= \frac{1}{\sqrt{2\pi t}^l} \int_{V(0)} e^{-\frac{|\phi(\xi)|^2}{2t}} f(\xi) \sqrt{\det g^L}(\xi) d\xi + O(t^{3/2}) \\ &= \mathcal{G}_t \left(Q \left([e^{-\frac{|\phi(\xi)|^2 - |\xi|^2}{2t}}] [\sqrt{\det g^L}] [f] \right) \right) + O(t^{3/2}) \\ &= \mathcal{G}_t \left(Q \left(1 + \frac{\delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v}(0) \xi^a \xi^b \xi^u \xi^v}{24t} \right) \left(1 + \frac{1}{6} R_{iuvu}(0) \xi^u \xi^v \right) [f] \right) \\ &\quad + O(t^{3/2}), \end{aligned}$$

since by Lemma 8 we have

$$\sqrt{\det g}(\xi) = e^{\frac{1}{2} \text{tr} \log g(\xi)} = 1 + \frac{1}{6} R_{iuvu}(0) \xi^u \xi^v + O(|\xi|^3) \quad (38)$$

and the error $O(|\xi|^3)$ leads to an error term $O(t^{3/2})$ in the Gaussian integral. Now

$$[f] = f(0) + \frac{\partial f}{\partial \xi^s}(0) \xi^s + \frac{1}{2} \frac{\partial^2 f}{\partial \xi^u \partial \xi^v}(0) \xi^v \xi^u.$$

But since the other two factor only contain monomials of even degree, all contributions containing $\partial f / \partial \xi^s$ are annihilated by Q . Thus

$$\begin{aligned} S(t)f(x) &= \mathcal{G}_t \left(Q \left(1 + \frac{\delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v}(0) \xi^a \xi^b \xi^u \xi^v}{24t} + \frac{1}{6} R_{iuvu}^L(0) \xi^u \xi^v \right) \right) f(0) \\ &\quad + \mathcal{G}_t \left(Q \left(\frac{1}{2} \frac{\partial^2 f}{\partial \xi^u \partial \xi^v}(0) \xi^v \xi^u \right) \right) + O(t^{3/2}) \end{aligned}$$

$$\begin{aligned}
&= \mathcal{G}_t \left(1 + \frac{1}{6} R_{iuv}^L(0) \xi^u \xi^v \right) f(0) + \frac{1}{2} \frac{\partial^2 f}{\partial \xi^u \partial \xi^v}(0) \mathcal{G}_t(\xi^v \xi^u) \\
&\quad + \mathcal{G}_t \left(\frac{\delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\alpha}{\partial \xi^u \partial \xi^v}(0) \xi^a \xi^b \xi^u \xi^v}{24t} \right) f(0) + O(t^{3/2}) \\
&= e^{-\frac{t R_{iuv}^L}{6}} f(0) + \frac{1}{24t} \mathcal{G}_t \left(\delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v}(0) \xi^a \xi^b \xi^u \xi^v \right) f(0) \\
&\quad + \frac{t}{2} \Delta f(0) + O(t^{3/2}).
\end{aligned}$$

The second term remains to be computed. By

$$\Delta \Delta \xi^a \xi^b \xi^u \xi^v = 8(\delta_{uv} \delta_{ab} + \delta_{ua} \delta_{vb} + \delta_{ub} \delta_{va}) \quad (39)$$

we obtain, having in mind that the *tension vector field* is the trace of the second fundamental form (see (Jost, 1998), (8.1.16), p.319 for the corresponding formula in local coordinates)

$$\begin{aligned}
&\frac{1}{24t} \mathcal{G}_t \left(\delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v}(0) \xi^a \xi^b \xi^u \xi^v \right) \\
&= \frac{t}{24} \delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^u \partial \xi^v}(0) (\delta_{ab} \delta_{uv} + \delta_{av} \delta_{ub} + \delta_{au} \delta_{vb}) \\
&= \frac{t}{24} \delta_{\alpha\beta} \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^a} \frac{\partial^2 \phi^\beta}{\partial \xi^b \partial \xi^b} + 2 \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^a \partial \xi^b} \right)(0) \\
&= \frac{t}{12} \delta_{\alpha\beta} \left(\frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^b} \frac{\partial^2 \phi^\beta}{\partial \xi^a \partial \xi^b} - \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^a} \frac{\partial^2 \phi^\beta}{\partial \xi^b \partial \xi^b} \right)(0) \\
&\quad + \frac{t}{8} \delta_{\alpha\beta} \frac{\partial^2 \phi^\alpha}{\partial \xi^a \partial \xi^a} \frac{\partial^2 \phi^\beta}{\partial \xi^b \partial \xi^b}(0) \\
&= \frac{t}{12} (\langle \mathbf{H}_\phi(e_a, e_b), \mathbf{H}_\phi(e_a, e_b) \rangle - \langle \mathbf{H}_\phi(e_a, e_a), \mathbf{H}_\phi(e_b, e_b) \rangle)(x) \\
&\quad + \frac{t}{8} |\tau(\phi)|^2(x)
\end{aligned}$$

where the vectors $e_{a,a=1,\dots,l}$ form an orthonormal base of $\phi_* T_x L$. But by the Gauss equations (see (Jost, 1998), Thm. 3.6.2, (3.6.7), p.151) we have

$$\begin{aligned}
&\langle \mathbf{H}_\phi(e_a, e_b), \mathbf{H}_\phi(e_a, e_b) \rangle - \langle \mathbf{H}_\phi(e_a, e_a), \mathbf{H}_\phi(e_b, e_b) \rangle \\
&= \langle R^M(e_b, e_a) e_a, e_b \rangle - \langle R^L(e_b, e_a) e_a, e_b \rangle.
\end{aligned}$$

Summing over a, b yields the statement. ■

Now, since

$$e^{\frac{t}{2}\Delta_L}f(x) = f(x) + \frac{t}{2}\Delta_L f(x) + O(t^{3/2})$$

Proposition 7 shows

$$\begin{aligned} \|B(t)f - e^{t/2\Delta_L}f\| &= \|e^{-tD}S(t)f - e^{t/2\Delta_L}f\| \\ &= \|(e^{-tD} - 1)\frac{t}{2}\Delta_L f\| + O(t^{3/2}) = O(t^{3/2}) \end{aligned}$$

for functions $f \in C^3(L)$. By Lemma 1 the family $(B(t))$ is Chernoff equivalent to the heat semigroup. This completes the proof of the Theorem.

Remark. According to Proposition 4 the normalized version $(T(t))$ (see (22)) of the family $(S(t))$ defined in (30) is also Chernoff equivalent to the heat semigroup.

5.5. PINNING THE HEAT KERNEL

In the case where M is Euclidean space the kernels in the integrals of section 5.3 are not only pseudo-Gaussian but truly Gaussian, i.e. they are induced by restricting the transition kernel of the Brownian motion in the ambient space to L . This is no longer true for arbitrary M . In order to obtain an analogous interpretation in the general case as well, we are now interested in the family obtained by using the *heat kernel* of the ambient manifold instead of the pseudo-Gaussian one. Let thus $p_t^M(x, y)$ be the heat kernel on M and consider on $C(L)$ the operator family

$$P_t f(x) := \sqrt{2\pi t}^{m-l} \int_L p_t^M(x, y) f(y) \text{vol}_L(dy). \quad (40)$$

It turns out that the calculation of an asymptotic formula for (40) is an application of the previous result as long as we take for granted the *Minakshisundaram-Pleijel expansion* (Minakshisundaram and Pleijel, 1949) of the heat kernel:

THEOREM 5. *Let $p_t^M(0, \xi)$ be the heat kernel on M in a normal coordinate neighbourhood around $x \in M$. Then we have the following asymptotic expansion*

$$\sqrt{2\pi t}^m e^{\frac{d_M(0, \xi)^2}{2t}} p_t^M(0, \xi) = \det^{-1/4} g^M(\xi) + \frac{t}{12} \text{Scal}_M(0) + O(t^{3/2}). \quad (41)$$

Proof: For a proof see (Roe, 1988), Prop. 5.25, p. 73 together with the computation of u_0 and u_1 on page 78. Note that we consider $\exp(t\Delta_M/2)$ instead of $\exp(t\Delta_M)$. \blacksquare

Thus we have in the same local coordinates as in Section 5.3

$$P_t f(x) = S(t) \left((\det^{-1/4} g^M(\xi) + \frac{t}{12} \text{Scal}_M(0)) f + O(t^{3/2}) \right),$$

where $S(t)$ is the operator in Proposition 7. By the expansion (8) for the metric tensor we obtain

$$\det^{-1/4} g^M(\xi) = 1 + \frac{1}{12} \text{Ric}_{uv}^M(0) \phi^u(\xi) \phi^v(\xi) + O(|\xi|^3), \quad (42)$$

where Ric^M denotes the *Ricci-tensor* of M . Therefore

$$\begin{aligned} \Delta \det^{-1/4} g^M(0) &= \sum_{i=1}^l \frac{\text{Ric}_{uv}^M}{6} \frac{\partial \phi^u}{\partial \xi^i} \frac{\partial \phi^v}{\partial \xi^i}(0) = \sum_{i=1}^l \langle e_i, \frac{\text{Ric}^M}{6} e_i \rangle(x) \\ &= \frac{1}{6} \overline{\text{Ric}}_{M/L}(x), \end{aligned} \quad (43)$$

where $\overline{\text{Ric}}_{M/L}$ denotes the partial trace of the Ricci-tensor using an orthonormal base e_i of $\phi_* T_x L$. Evaluating (42) at $\xi = 0$ we obtain

$$\begin{aligned} P_t f(x) &= e^{t \left(-\frac{\text{Scal}_L}{4} + \frac{|\tau|^2}{8} + \frac{\overline{R}_{M/L}}{12} \right)(x)} \left(1 + \frac{t}{12} \text{Scal}_M(x) \right) f(x) \\ &\quad + \frac{t}{2} \Delta_L (\det^{-1/4} g^M f)(x) + O(t^{3/2}) \\ &= e^{-\frac{t \text{Scal}_L}{4} + \frac{t |\tau|^2}{8} + \frac{t (\overline{R}_{M/L} + \text{Scal}_M)}{12}} f(x) \\ &\quad + \frac{t}{2} (\Delta_L + \Delta_L \det^{-1/4} g^M(x)) f(x) + O(t^{3/2}). \end{aligned}$$

and with (43) we finally arrive at the asymptotic formula

$$\begin{aligned} P_t f(x) &= e^{t \left(-\frac{\text{Scal}_L}{4} + \frac{|\tau|^2}{8} + \frac{\overline{R}_{M/L} + \overline{\text{Ric}}_{M/L} + \text{Scal}_M}{12} \right)(x)} f(x) + \frac{t}{2} \Delta_L f(x) + O(t^{3/2}). \end{aligned} \quad (44)$$

Therefore the analogues of Theorem 4 and the Remark at the end of the previous section consist of the following statement:

THEOREM 6. *Using the notations from above the operator families $(B(t))$ and $(T(t))$ defined on $C(L)$ by*

$$B(t)f(x) := \frac{e^{t \left(\frac{\text{Scal}_L}{4} - \frac{|\tau_\phi|^2}{8} - \frac{\overline{R}_{M/L} + \overline{\text{Ric}}_{M/L} + \text{Scal}_M}{12} \right)(x)}}{\sqrt{2\pi t}^{l-m}} \int_L p_t^M(x, y) f(y) \text{vol}_L(dy)$$

and

$$T(t)f(x) = \frac{\int_L p_t^M(x, y)f(y)\text{vol}_L(dy)}{\int_L p_t^M(x, y)\text{vol}_L(dy)}$$

are both Chernoff equivalent to the heat semigroup on L .

6. Limit Densities for the Pinning Construction under global normalization

6.1. INTRODUCTION

We return to the introduction of section 5. The operator families $S = (S(t))$ defined by (21) are not proper in general. In our examples their rescaled versions $T = (T(t))$ and $B = (B(t))$, cf. (22) and (24)), become proper and Chernoff equivalent to the heat semigroup and thus, applying the pinning construction with either T or B yields (assuming an appropriate choice of the interpolation family \mathcal{Q}) weak convergence of the measures to the law \mathbb{W}_L^x of Brownian motion on L .

On the other hand we can perform the pinning construction directly with S . Then for each partition \mathcal{P} of $[0, 1]$ the comparison of the induced measures on $L^\mathcal{P}$ corresponding to B and S yields

$$\frac{d\mathbb{P}_{L, \mathcal{P}, S}^x}{d\mathbb{P}_{L, \mathcal{P}, B}^x}(x; y_1, \dots, y_r) = e^{(t_1 - t_0)D(x) + \sum_{k=1}^{r-1} (t_{k+1} - t_k)D(y_k)}.$$

where we use the symbol of the family as an additional index. If we choose the same interpolation family for both B and S this implies

$$\frac{d\mathbb{P}_{\mathcal{P}, S}^x}{d\mathbb{P}_{\mathcal{P}, B}^x}(\omega) = e^{(t_1 - t_0)D(x) + \sum_{k=1}^{r-1} (t_{k+1} - t_k)D(\omega(t_k))}.$$

on the path space $C_M[0, 1]$. In the exponent, we have Riemann sums which converge uniformly on compact subsets of $C_M[0, 1]$ to the corresponding integrals. According to Theorem 3 for $|\mathcal{P}_k| \rightarrow 0$ the sequence $(\mathbb{P}_{\mathcal{P}_k, B}^x)$ is uniformly tight. Therefore the measures $\mathbb{P}_{\mathcal{P}_k, S}^x$ converge weakly over the space $C_M[0, 1]$ to a measure ν_L^x which is concentrated on $C_L[0, 1]$ with density

$$\frac{d\nu_L^x}{d\mathbb{W}_L^x}(\omega) = \lim_{|\mathcal{P}| \rightarrow 0} \frac{d\mathbb{P}_{\mathcal{P}, S}^x}{d\mathbb{P}_{\mathcal{P}, B}^x}(\omega) = \exp\left(\int_0^1 D(\omega(s))ds\right). \quad (45)$$

As a first application we get for $M = L$ a result which for geodesic interpolation on L was essentially contained with a completely different proof first in (Andersson and Driver, 1999).

COROLLARY 7. *Let in the construction above q be the pseudo-Gaussian kernel on L*

$$q_t(x, y) := \frac{1}{(2\pi t)^{l/2}} e^{-\frac{d_L(x, y)^2}{2t}}.$$

Let \mathcal{Q} be an arbitrary interpolating family such that the sequence $(\mathbb{P}_{\mathcal{P}_k}^x)$ is tight over $C_L[0, 1]$. Then the limit measure ν_L^x corresponding to these kernels under the pinning construction is equivalent to Wiener measure \mathbb{W}_L^x on L with density

$$\frac{d\nu_L^x}{d\mathbb{W}_L^x}(\omega) = \exp\left(\frac{1}{6} \int_0^1 \text{Scal}(\omega(s)) ds\right).$$

We can state two more density results in the spirit of Corollary 7. Namely, if we use the following pseudo-Gaussian density

$$q_t(x, y) := \frac{1}{(2\pi t)^{l/2}} e^{-\frac{d_M(x, y)^2}{2t}} \quad (46)$$

but integrate with respect to the volume form on L we obtain the following result from Theorem 4 in the same way as in the preceding proof.

COROLLARY 8. *The family of measures obtained by the pinning construction to the rescaled restriction (46) of the pseudo-Gaussian kernel on M integrated with respect to L converges weakly to a measure ν_L^x which is equivalent to Wiener measure on L with density*

$$\frac{d\nu_L^x}{d\mathbb{W}_L^x}(\omega) = \exp\left(\int_0^1 \left\{ \frac{1}{4} \text{Scal}_L - \frac{1}{8} |\tau|^2 - \frac{1}{12} \bar{R}_{M/L} \right\} (\omega(s)) ds\right).$$

If we use the properly normalized heat kernel on M , namely

$$q_t(x, y) := (2\pi t)^{(m-l)/2} p_t^M(x, y) \quad (47)$$

we get from Theorem 6:

COROLLARY 9. *The family of measures obtained by the pinning construction applied to the restriction of the rescaled heat kernel (47) of M to L converges weakly to a measure ν_L^x which is equivalent to Wiener measure on L with density*

$$\frac{d\nu_L^x}{d\mathbb{W}_L^x}(\omega) = \exp\left(\int_0^1 D(\omega(s)) ds\right)$$

where the function D on L is defined by

$$D(y) = \left(\frac{1}{4} \text{Scal}_L - \frac{1}{8} |\tau_\phi|^2 - \frac{1}{12} (\bar{R}_{M/L} + \bar{\text{Ric}}_{M/L} + \text{Scal}_M) \right) (y). \quad (48)$$

Note that the function D in (48) is precisely the function in the exponent in (2) in Theorem 1. The explanation will be given below.

6.2. THE PINNING CONSTRUCTION AS CONDITIONAL PROBABILITY

We now want to apply our convergence results to the comparison of the Wiener measures on the manifolds L and M respectively. To do so, we compare the pinning construction applied to the family S with another one obtained by normalizing at all partition times simultaneously. To be precise we construct from $\mathbb{P}_{L,\mathcal{P},S}^x$ the probability measure

$$\mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}(dy_1, \dots, dy_r) := C_{\mathcal{P}}(x) \mathbb{P}_{L,\mathcal{P},S}^x(dy_1, \dots, dy_r),$$

where the constant is chosen so that this becomes a probability distribution. In the special case where, similarly as in Corollary 9, the kernel is given by the restriction of the heat kernel (p_t^M) on M , the resulting probability measure $\mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}$ on $L^{\mathcal{P}}$ describes the *marginal distribution of the M -Brownian motion W_M^x under the condition that it visits L at all times t_i in the partition*. (Note that now the rescaling constants in (47) are no longer necessary because they are incorporated in the new constant $C_{\mathcal{P}}(x)$.) Since the volume measure of L can be obtained from the volume measure on M by the usual surface measure limiting procedure, we can compute the measure $\mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}$ also by

$$\begin{aligned} & \mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}(A_1 \times \dots \times A_r) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\int_{U_{\varepsilon}(A_1) \times \dots \times U_{\varepsilon}(A_r)} p_{t_1}^M(x, dy_1) \dots p_{1-t_{r-1}}^M(y_{r-1}, dy_r)}{\int_{U_{\varepsilon}(L)^r} p_{t_1}^M(x, dy_1) \dots p_{1-t_{r-1}}^M(y_{r-1}, dy_r)}. \end{aligned} \tag{49}$$

Thus we can consider $\mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}$ as a kind of surface measure induced by the heat kernel of M on $L^{\mathcal{P}}$. Hence the upper index ' Σ '. To point out the difference to the pinning distribution marginal we write the latter as

$$\mathbb{P}_{L,\mathcal{P}}^x(A_1 \times \dots \times A_r) = \lim_{\varepsilon \rightarrow 0} \frac{\int_{U_{\varepsilon}(A_1) \times \dots \times U_{\varepsilon}(A_r)} p_{t_1}^M(x, dy_1) \dots p_{1-t_{r-1}}^M(y_{r-1}, dy_r)}{\prod_{k=1}^r \int_{U_{\varepsilon}(L)} p_{t_k-t_{k-1}}^M(y_{k-1}, dz)}$$

letting $y_0 = x$, the difference being that in this last expression the kernel is renormalized at each partition time.

Let $\mathbb{W}_{\mathcal{P}}^{x,\Sigma}$ denote the law on $C_M[0,1]$ which we get from $\mathbb{P}_{L,\mathcal{P}}^{x,\Sigma}$ by using *M -Brownian bridge measures* for the interpolating family \mathcal{Q} . Then $\mathbb{W}_{\mathcal{P}}^{x,\Sigma}$ can be considered as the law of *M -Brownian motion* conditioned to be on the submanifold L at the times $t_k \in \mathcal{P}$. Since the normalizing map $\nu \mapsto \frac{1}{\nu(C_M[0,1])}\nu$ is continuous with respect to the topology of 'weak convergence' on the cone of finite positive measures and since the

convergence of the nonnormalized measures is known from Corollary 9 we get finally the following reformulation of Theorem 1: Let $x \in L$. As the mesh $|\mathcal{P}|$ of a partition \mathcal{P} of $[0, 1]$ converges to 0, the conditional law

$$\mathbb{W}_M^x(d\omega | \omega(t_k) \in L, t_k \in \mathcal{P})$$

of Brownian motion on M , conditioned to visit L at all partition times, tends weakly over $C_M[0, 1]$ to the measure μ_L^x which is equivalent to \mathbb{W}_L^x with density (2).

7. Conclusion

Let L be a smooth closed Riemannian manifold. For various one-parameter families $S = (S(t))$ of kernel operators of the form

$$S(t)f(x) = \int_L q(t, x, y)f(y)vol_L(dy) \quad (50)$$

we can verify the short time asymptotics

$$S(t)f(x) = e^{tD(x)} + \frac{t}{2}\Delta_L f(x) + O(t^{3/2}) \quad (51)$$

for all $f \in C^3(L)$ and some function $D \in C(L)$ which depends on S and is a combination of curvature terms.

For each partition \mathcal{P} of $[0, 1]$ and interpolation by geodesics or Brownian bridges the family S induces a measure on the path space $C_L[0, 1]$. As the partition gets finer we prove convergence in law of these measures to a measure ν_L^x which is equivalent to the law \mathbb{W}_L^x of L -valued Brownian motion with Radon-Nikodym density

$$\frac{d\nu_L^x}{d\mathbb{W}_L^x}(\omega) = \exp\left(\int_0^1 D(\omega(s))ds\right). \quad (52)$$

If $q(t, x, y)$ is replaced by $\tilde{q}(t, x, y) = \frac{q(t, s, y)}{\int_L q(t, x, y)vol_L(dy)}$ then the associated probability measures on the path space converge to \mathbb{W}_L^x .

If $q(t, x, y)$ is the restriction of the heat kernel on a surrounding manifold M to L then the above results imply the convergence of the conditional law of Brownian motion on M , conditioned to return to L at all partition times to a measure of the form (52) with an explicitly known function D . Here ν_L^x can be viewed as the infinite dimensional surface measure induced by \mathbb{W}_M^x on the set $C_L[0, 1] \subset C_M[0, 1]$.

The key tools for these results are a) asymptotic computation of Gaussian integrals combined with b) various differential geometric computations for the proof of (51) and c) a new version of Chernoff's

theorem in semigroup theory combined with d) tightness results for the convergence of the measures.

References

L. Andersson and B.K. Driver. Finite dimensional approximations to Wiener measure and path integral formulas on manifolds. *J. Funct. Anal.*, 165:430–498, 1999.

M.P. doCarmo. *Riemannian Geometry*. Birkhäuser, Boston, 1992.

R.M. Dudley. *Real Analysis and Probability*. Wadsworth and Brooks/Cole, Pacific Grove, California, 1989.

S.N. Ethier and T.G. Kurtz. *Markov Processes, Characterizations and Convergence*. John Wiley & Sons, New York etc., 1986.

J. Jost. *Riemannian Geometry and Geometric Analysis, 2nd edition*. Springer, Heidelberg-New York, 1998.

T. Kato. *Perturbation Theory for Linear Operators, 2nd. ed.* Springer, New York, 1980.

S. Minakshisundaram and A. Pleijel. Some Properties of the Eigenfunctions of the Laplace operator on Riemannian Manifolds. *Canad. J. Math.*, 1:242–256, 1949.

J. Roe. *Elliptic Operators, topology and asymptotic methods*. Longman, London, 1988.

N. Sidorova. The Smolyanov surface measure on trajectories in a Riemannian manifold. Infinite Dimensional Analysis, Quantum Probability and Related Topics, 7,461-471, 2004.

N. Sidorova, O.G. Smolyanov, H.v. Weizsäcker, and O. Wittich. Conditioning Brownian Motion to Small Tubular Neighborhoods. *in preparation*, 2005.

N. Sidorova, O.G. Smolyanov, H.v. Weizsäcker, and O. Wittich. The Surface Limit of Brownian Motion in Tubular Neighbourhoods of an embedded Riemannian Manifold. *J. Funct. Anal.*, 206:391–413, 2004.

O.G. Smolyanov, H.v. Weizsäcker, and O. Wittich. Brownian motion on a manifold as limit of stepwise conditioned standard Brownian motions. In *Stochastic processes, Physics and Geometry: New Interplays. II: A Volume in Honour of S. Albeverio*, volume 29 of *Can. Math. Soc. Conf. Proc.*, pages 589–602. Am. Math. Soc., 2000.

O.G. Smolyanov, H.v. Weizsäcker, and O. Wittich. Chernoff’s theorem and the construction of Semigroups. In *Evolution Equations: Applications to Physics, Industry, Life sciences and Economics - EVEQ 2000*, M. Ianelli, G. Lumer (eds.), pages 355–364. Birkhäuser, 2003.

D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer, New York, 1979.

A.G. Tokarev. Unpublished Notes. Moscow, 2001.

O. Wittich. An Explicit Local Uniform Large Deviation Bound for Brownian Bridges. *Stat. Prob. Lett.*, 73:51–56, 2005.

O. Wittich. Effective Dynamics on Small Tubular Neighbourhoods. *in preparation*, 2005.

